A Description of elicited data sets

Prosodic data and analyses used for this paper are available as on-line supplementary material at the following webpage: `http://www.krisyu.org/blog/supp-material-invariability-samoan-interface.html`.

A.1 *Speech rate*

This data set was elicited from our primary consultant in Los Angeles to study the realization of the absolutive high under different speech rates and under tonal crowding, where the number of syllables between the absolutive high and neighboring stress was varied. The consultant was asked to produce the sentences below, as well as their intransitive counterparts (using the verb [manoŋi] 'to be smelly' (in randomized order), first at a comfortable pace, and then a fast pace, and then a slow pace. One fluent repetition of each sentence was recorded.

(26) Transitive sentences, 3 syllable subject
   a. na lagona e malini le liona i le aoauli. 'The marines heard the lion in the afternoon.'
   b. na lagona e malini liona i le aoauli. 'The marines heard the lions in the afternoon.'
   c. na lagona e malini manu i le afiafi. 'The marines heard the birds in the evening.'
   d. na lagona e malini le manu i le afiafi. 'The marines heard the bird in the evening.'
   e. na lagona e malini le mea i le afiafi. 'The marines heard the thing in the evening.'
   f. na lagona e malini mea i le taeao. 'The marines heard the thing in the morning.'
   g. na lagona e malini iʔa i le taeao. 'The marines heard the fish in the morning.'
   h. na lagona e malini ia i le taeao. 'The marines heard him in the morning.'

(27) Transitive sentences, 2 syllable CVCV subject
   a. na lagona e manu le liona i le aoauli. 'The birds heard the lion in the afternoon.'
   b. na lagona e manu liona i le aoauli. 'The birds heard the lions in the afternoon.'
   c. na lagona e manu le manu i le afiafi. 'The birds heard the bird in the evening.'
   d. na lagona e manu manu i le afiafi. 'The birds heard the birds in the evening.'
   e. na lagona e manu le mea i le taeao. 'The birds heard the thing in the evening.'
   f. na lagona e manu mea i le taeao. 'The birds heard the things in the morning.'
   g. na lagona e manu iʔa i le taeao. 'The birds heard the fish in the morning.'
   h. na lagona e manu ia i le taeao. 'The birds heard him in the morning.'

(28) Transitive sentences, 2 syllable CVV subject
   a. na lagona e moa le liona i le aoauli. 'The chickens heard the lion in the afternoon.'
   b. na lagona e moa liona i le aoauli. 'The chickens heard the lions in the afternoon.'
   c. na lagona e moa manu i le afiafi. 'The chickens heard the birds in the evening.'
   d. na lagona e moa le manu i le afiafi. 'The chickens heard the bird in the evening.'
   e. na lagona e moa mea i le afiafi. 'The chickens heard the things in the evening.'
   f. na lagona e moa le mea i le afiafi. 'The chickens heard the thing in the evening.'
   g. na lagona e moa iʔa i le taeao. 'The chickens heard the fish in the morning.'
   h. na lagona e moa ia i le taeao. 'The chickens heard him in the morning.'

(29) Transitive sentences, 1 syllable CV: subject
   a. na lagona e lā le liona i le aoauli. 'Sun heard the lion in the afternoon.'
   b. na lagona e lā liona i le aoauli. 'Sun heard the lions in the afternoon.'

c. na lagona e lā manu i le afiafi. ‘Sun heard the birds in the evening.’
d. na lagona e lā le manu i le afiafi. ‘Sun heard the bird in the evening.’
e. na lagona e lā mea i le afiafi. ‘Sun heard the things in the evening.’
f. na lagona e lā le mea i le afiafi. ‘Sun heard the thing in the evening.’
g. na lagona e lā iʔa i le taeao. ‘Sun heard the fish in the morning.’
h. na lagona e lā ia i le taeao. ‘Sun heard him in the morning.’

(30) Transitive sentences, VV subject (pronoun)
a. na lagona e ia le liona i le aoauli. ‘He heard the lion in the afternoon.’
b. na lagona e ia liona i le aoauli. ‘He heard the lions in the afternoon.’
c. na lagona e ia manu i le afiafi. ‘He heard the birds in the evening.’
d. na lagona e ia le manu i le afiafi. ‘He heard the bird in the evening.’
e. na lagona e ia mea i le afiafi. ‘He heard the things in the evening.’
f. na lagona e ia le mea i le afiafi. ‘He heard the thing in the evening.’
g. na lagona e ia iʔa i le taeao. ‘He heard the fish in the morning.’
h. na lagona e ia ia i le taeao. ‘He heard him in the morning.’

A.2 *Tautala lelei*

This data set was elicited with 2 consultants (f03, f05) in Auckland, New Zealand in July 2015 and was designed to check for the presence of the absolutive high under different discourse conditions. Four sets of question-answer pairs were elicited: Two with transitive verbs ([lalaŋa] ‘weave’, taking an inanimate object; [laŋona] ‘hear’, taking an animate object), and two with intransitive verbs ([malaŋa] ‘journey’, taking an inanimate PP object; [leaŋa] ‘be bad’, taking an animate PP object).

In (31) below for the [lalaŋa] set, we show the different question types used to generate different discourse conditions and the different answer types elicited. We recorded any question-answer pairs that the consultants accepted. For this paper, the data set from f03 includes only the two sets with inanimate objects.

(31) Question types
a. wh focus on subject

o ai na lalaŋa-ina le mamanu i le aso:?
TOP wh PAST weave-INA DET.SG design OBL DET.SG day

‘Who wove the design today?

b. wh focus on object

o le a: na fai e le malini i le aso:?
TOP DET.SG wh PAST do ERG DET.SG marine OBL DET.SG day

‘What did the marine do today?

c. wh focus on VP

o le a: le mea na lalaga-ina e le malini i le aso:?
TOP DET.SG wh DET.SG thing PAST weave-INA ERG DET.SG marine OBL DET.SG day

‘What did the marine weave today?

d. broad (polarity) focus

na lalaga e le malini le mamanu i le aso:?
PAST weave ERG DET.SG marine DET.SG design OBL DET.SG day

‘Did the marine weave the design today?

e. corrective focus on subject

na lalaga e le pailate le mamanu i le aso:.
PAST weave ERG DET.SG pilot DET.SG design OBL DET.SG day

'The pilot wove the design today.'

f. corrective focus on object

na lalaga e le malini le ato i le aso:.
PAST weave ERG DET.SG marine DET.SG basket OBL DET.SG day

'The pilot wove the basket today.'

(32) Answer types

a. VSO

(leai,) na lalaga e le malini le mamanu i le aso:.
(no) PAST weave ERG DET.SG marine DET.SG design OBL DET.SG day

'(No,) The marine wove the design today.'

b. VSO (negative polarity)

leai, e le?i lalaga-ina e le malini le mamanu i le aso:.
no, pres neg weave-INA ERG DET.SG marine DET.SG design OBL DET.SG day

'No, it's not the case that the marine wove the design today.'

c. VOS

na lalaga-ina le mamanu e le malini i le aso:.
PAST weave-INA DET.SG design ERG DET.SG marine OBL DET.SG day

'The design was woven by the marine today.'

d. fronted subject

(leai,) o le malini na lalaga-ina le mamanu i le aso:
(no,) TOP DET.SG marine PAST weave-INA DET.SG design OBL DET.SG day

'(No,) It was the marine that wove the design today.'

e. fronted object

(leai,) o le mamanu na lalaga-ina e le malini i le aso:
(no,) TOP DET.SG design PAST weave-INA ERG DET.SG marine OBL DET.SG day

'(No,) it was the design that the marine wove today.'

A.3 *Tautala leaga*

This data set was elicited with 2 consultants (f03, f05) in Auckland, NZ in July 2015. For each question in the *tautala lelei* data set, consultants were asked: Out of the responses they found permissible, which was the one they most preferred? A selection of these preferred responses were then recorded in *tautala leaga*. Both consultants preferred to read sentences written in *tautala lelei* but convert them on the fly into *tautala leaga*, saying this was a very natural thing for them to do. Speaker f03 explicitly said she was dropping the ergative *e*, but this was not the case for f05; both speakers left in the oblique *i*. For f03, only two sets from the *tautala lelei* data were included—the two sets with inanimate objects ([lalaŋa] 'weave' and [malaŋa] 'journey'. Some of the speakers' preferred responses included the particle *ia*, which is beyond the scope of this paper—these were excluded from the data set for this paper.

The data set included for f03 is:

Set	Discourse context	Answer type
lalaŋa	broad/polarity	VSO
	wh focus, object	fronted object
	corrective focus, subject	fronted subject
	corrective focus, object	fronted object
malaŋa	wh focus, subject	fronted subject
	broad/polarity	V-S-PP
	wh focus, object	V-S-PP
	corrective focus, subject	fronted subject
	corrective focus, object	VSO

The data set included for f05 is:

Set	Discourse context	Answer type
lalaŋa	wh focus, subject	fronted subject
	broad/polarity	VSO
	wh focus, VP	VSO
laŋona	wh focus, subject	VSO
	broad/polarity	VSO
	wh focus, VP	VSO
	corrective focus, subject	VSO
	corrective focus, object	VOS
malaŋa	wh focus, subject	fronted subject
	broad/polarity	V-S-PP
	wh focus, VP	V-S-PP
	wh focus, object	V-S-PP
leaŋa	wh focus, subject	fronted subject
	wh focus, VP	V-S-PP
	wh focus, object	fronted object
	corrective focus, subject	fronted subject
	corrective focus, object	fronted object

A.4 Basic coordination

This data set was elicited with 1 consultant (f03) in Auckland, NZ in July 2015 and was designed to check for the presence of a high tone in a range of coordination structures. All sentences were elicited under "broad focus" as answers with negative polarity, as described in A.2. The following sentences were used to elicit the conjunction *ma*, their analogues were also used to elicit the disjunctive coordinator [poʔo]/[peː] ([peː] for disjunction of verbs):

(33) Transitives, coordinated subjects

a. leai, e leʔi lalaŋa Manoŋi ma Malini ato.

'No, it's not the case that Manogi and Malini wove the baskets.'

b. leai, e leʔi laŋona e Manoŋi ma Malini maile.

'No, it's not the case that Manogi and Malini heard the dogs.'

c. leai, e leʔi momoli e Manoŋi ma Malini mamanu.

'No, it's not the case that Manogi and Malini dropped off the designs.'

(34) Intransitives, coordinated subjects

a. leai, e leʔi leaŋa Manoŋi ma Malini i le maile.

'No, it's not the case that Manogi and Malini were bad to the dog.'

b. leai, e leʔi galue Manoŋi ma Malini i le mamanu.

'No, it's not the case that Manogi and Malini worked on the design.'

c. leai, e leʔi malaŋa Manoŋi ma Malini i le moana.

'No, it's not the case that Manogi and Malini journeyed to the sea.'

(35) Transitives, coordinated objects

a. leai, e leʔi lalaŋa e Malini mamanu ma le ato.

'No, it's not the case that Malini wove the designs and the basket.'

b. leai, e leʔi laŋona e Malini liona ma le maile.

'No, it's not the case that Malini heard the lions and the dog.'

c. leai, e leʔi momoli e Malini meleni ma puligi.

'No, it's not the case that Malini dropped off the melons and the puligi (a Samoan spiced pudding).'

(36) Intransitives, coordinated objects

a. leai, e leʔi leaŋa Malini i le liona ma le maile.

'No, it's not the case that Malini was bad to the lion and the dog.'

b. leai, e leʔi galue Malini i le mamanu ma le ato.

'No, it's not the case that Malini worked on the design and the basket.'

c. leai, e leʔi malaŋa Malini i le moana ma le mauga.

'No, it's not the case that Malini journeyed to the sea and the mountain.'

(37) Coordinated verbs

a. leai, e leʔi lalaŋa ma momoli e Malini mamanu.

'No, it's not the case that Malini wove and dropped off the designs.'

b. leai, e leʔi ŋalue ma fufulu e Malini mamanu.

'No, it's not the case that Malini worked on and cleaned the designs.'

c. leai, e leʔi leaŋa ma auleaŋa le maile i le fale.

'No, it's not the case that the dog in the house is bad and ugly.'

d. leai, e leʔi galue Malini ma alu e? momoli mamanu.[1]

'No, it's not the case that Malini worked on and went to drop off the designs.'

A.5 Prepenultimate stress

This data set was elicited with 2 consultants (f03, f05) in Auckland, NZ in July 2015 and was designed to test if various morphosyntactically conditioned high tones were edge tones, rather than a modification of the pitch accent associated with stressed moras. Therefore, target items varied stress position by including English names that were familiar to the consultants to place stress initially: Initial stress (*Romeo*, *Melanie*), penult stress (*Marilla*, *Manogi*), final stress (*Gabrielle*, $Enel\bar{e}$). Then, minimally different sentences were designed to elicit the different morphosyntactically conditioned high tones. All items were designed to be under broad focus (with focus on polarity in the answers to questions), except in fronting examples (42). The set of sentences in 38 illustrates the range of sentences used for *Melanie*. For the conjunction, disjunction, and first name last name conditions, the intransitive verb *malaga* 'to journey, travel' was additionally used, for a total of 9 sentences per target name. The "first name last name" condition was designed as a baseline condition for when we didn't expect to see a high tone.

[1]The consultant was inconsistent here in whether she put in the ergative *e*. She didn't in conjunction, but did in disjunction.

(38) Sentences with transitive verbs for *Melanie* (initial stress)

a. Question

na momoli e Melanie ma Malu le liona? (conjunction)
PAST drop.off ERG Melanie CONJ Malu DET.SG lion

'Did Melanie and Malu drop off the lion?'

b. Answer

leai, e leʔ momoli e Melanie ma Malu le liona. (conjunction)
no, PRES NEG drop.off ERG Melanie conj Malu DET.SG lion

'No, it's not the case that Melanie and Malu dropped off the lion.'

(39) leai, e leʔ momoli e Melanie-Mamalu le liona? (first name, last name)

'No, it's not the case that Melanie-Mamalu dropped off the lion.'

(40) leai, e leʔ momoli e Melanie Mamalu i le liona? (absolutive)

'No, it's not the case that Melanie dropped off Mamalu with the lion.'

(41) leai, e leʔ momoli e Melanie poʔo Malu i le liona? (disjunction)

'No, it's not the case that Melanie or Malu dropped off the lion.'

(42) Fronting examples

a. Question

o ai na opo-ina Mamalu i le moega?
TOP wh PAST hug-INA Mamalu OBL DET.SG bed

'Who hugged Mamalu in bed?'

b. Answer

o Melanie na opo-ina Mamalu i le moega.
TOP wh PAST hug-INA Mamalu OBL DET.SG bed

'Melanie hugged Mamalu in bed.'

a. Question

o ai na lavea iaː Mamalu i le aʔoga.
TOP wh PAST hurt OBL Mamalu OBL DET.SG school

'Who was hurt by Mamalu at school?'

b. Answer

o Melanie na lavea iaː Mamalu i le aʔoga.
TOP wh PAST hurt OBL Mamalu OBL DET.SG school

'Melanie was hurt by Mamalu at school.'

B Formal mechanisms for parsing and spellout

A simple example grammar. Providing a new syntax for Samoan case is not among the goals of this paper. The present interest is the syntax/prosody interface, so to make our proposals about the interface more tangible and comprehensible, we have simply adopted a particular perspective on Samoan case based on Collins' studies of Samoan case and syntax (Collins, 2016, 2015, 2014). We expect that a similar account of the syntax/prosody interface could be provided with any of the prominent alternatives in the syntax literature. Collins' approach is fairly well worked out, and so it is fairly easy to support our claim in Section 7.1 that the structures he proposes could be defined by a version of minimalist grammar (MG). Here we flesh that argument out just slightly by presenting the beginnings of a formal MG for Samoan. This sketch is very far from complete and correct; and it does not implement many aspects of Samoan treated by Collins. It is just meant to indicate the basic mechanisms that MGs have available, in support of the claims of Section 7. The key thing is that the set of derivation trees is regular. The derivation trees are regular, in a standard sense explained below, not just for this extremely simple illustrative example MG, but for *any* MG, *no matter how complex* (Michaelis, 1998; Kobele et al., 2007). Consequently, given any reasonable proposals about how case marking happens, we will be able to encode that case marking either as a postsyntactic 'transduction' on the derivation or directly in the syntax as some other linguists have proposed.

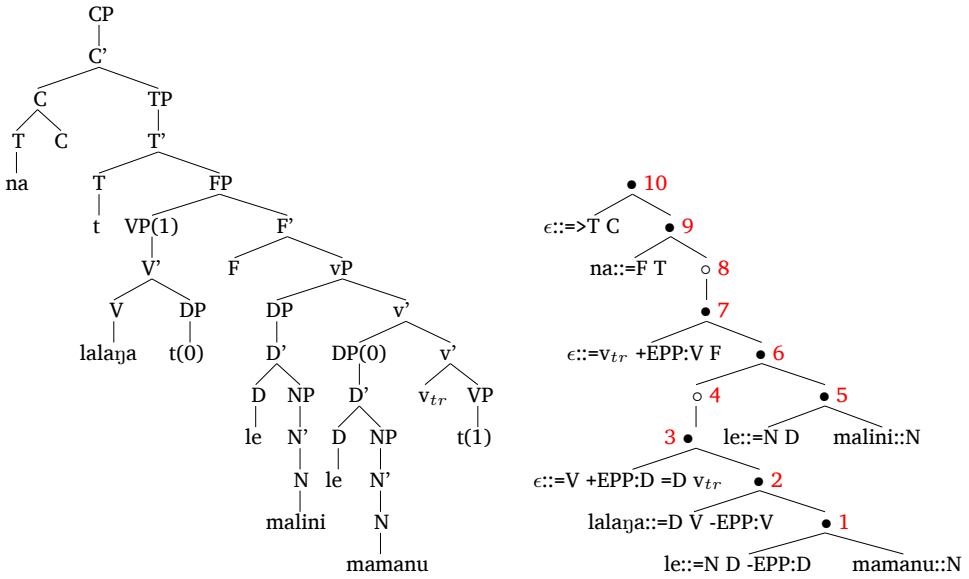
A minimalist grammar is given by its lexicon, where each lexical item specifies some phonological content (here we use the spelling) and a sequence of features. Here we introduce MG notation with only very brief explanation; see Stabler (2010) and references cited there for extensive, detailed treatments of all the mechanisms. But the basic idea is quite simple. The feature =D indicates that a DP is to be selected; the feature =>T indicates that a TP is to be selected but also triggers the T head to move to the head of the selecting category. The feature +wh triggers the movement of a +wh phrase to specifier; the feature +EPP:D triggers the movement of a -EPP:D phrase; and the feature +EPP:V triggers the movement of a -EPP:V phrase. We use (x) to indicate an optional feature x and {x,y} to indicate 'exactly one of x,y'.[2]

ϵ::=>T ({+top,+wh}) C	selects TP with head raising, optionally moves -top or -wh, to form CP
ϵ::=V +EPP:D =D v_{tr}	selects VP, moves EPP:D, selects DP, to form vp
ϵ::=V =D v	selects VP, selects DP, to form vp
ϵ::{=v,=v_{tr}} +EPP:V F (-f)	selects vp, moves -EPP:V, to form F, optionally moves to +f
na::=F (+f) T	'past', selects FP, optionally moves FP, to form T
lalaηa(-ina)::=D V -EPP:V	'weave', selects DP to form VP, then moves to EPP
ηalue::V -EPP:V	'work', a VP, moves to +EPP:V
le::=N D (-EPP:D)	selects NP to form DP, optionally moves to +EPP:D
malini::N	noun 'marine'
mamanu::N	noun 'design'

This lexicon derives, for example, the structure on the left below, in 10 steps numbered on the right, where (external) merge is indicated by •, and move (or 'internal merge') is indicated by ◦.

(43) Deriving example (2a) with structure (11a) before case marking:

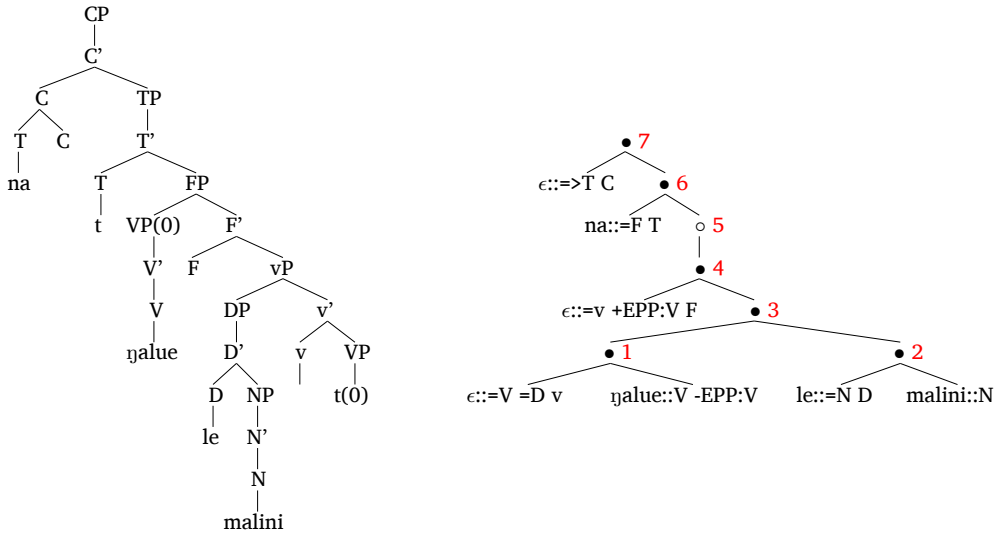
[2]The tiny lexicon below suffices for the derivations shown here, but it is obviously very much simplified from what is required. See Collins (2016, 2014) for a much more careful discussion. It appears that all the mechanisms proposed by Collins, including his 'conditional' and 'unconditional' movement triggers (2016, §7.2), could be added to MGs without affecting expressive power or any of the other basic computational properties that are relevant here.



Notice that the predicate initial structure is formed by first raising the object out of the VP (at step 4) leaving trace t(0), and then moving the VP containing that trace to spec,FP (at step 8) leaving trace t(1), as proposed by Collins (2016). We will do the case marking in the spellout, described below.

We see an example intransitive derivation here, in which we have numbered the merge steps.

(44) Deriving (2b) with structure (11b) before case marking:



In both derivations, each step involves checking and deleting a pair of features (though more complex feature checking regimes can be added easily, as discussed by Stabler 2010). In this last derivation, the 7 steps check the 14 lexical features to leave a completed CP, as follows:

step	result of feature checking	input to spellout
1	=D v, -EPP:V	(ϵ, ŋalue)
2	D	le malini
3	v, -EPP:V	(le malini, ŋalue)
4	+EPP:V F,-EPP:V	(le malini, ŋalue)
5	F	ŋalue le malini
6	T	na ŋalue le malini
7	C	na ŋalue le malini

Kobele et al. (2007) point out that this feature checking can be done by a finite state bottom-up tree automaton, so the set of derivation trees is 'regular', a finite state tree language.

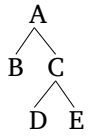
The methods of Section 3 extend straightforwardly to the additional constructions in Section 5. We add these lexical items to our example grammar:

laŋona::V -EPP:V	intransitive verb 'hear'
malaŋa::V -EPP:V	intransitive verb 'travel'
Malu::D	name
ma::=D =D D	coordinator 'and'
po'o::=D =D D	coordinator 'or'
ʔo::=D D -top	topicalizer

With these lexical items, we can derive (18) and (20).

Post-syntactic case and tone marking. We propose to implement case marking as something that happens in the spellout of structure, but we implement this in two steps. First we mark in the derivation tree where the case marking should happen, and then we spell out the structure in a way that departs only minimally from the standard spellout that is indicated, for example, in the simple derivations presented above.

In technical treatments of this kind of 'tree transducer', it is common to use the standard 'term notation' for trees. That is, instead of drawing a tree like this:



we can write the same structure as a term A(B,C(D,E)). To define tree transducers, we allow these terms to contain variables ranging over subtrees. For an introduction to tree trandsucers, see for example Comon et al. (2007, §6.3) or Graf (2013, §4.1.2).

A bottom-up identity tree transducer for the derivation tree language traverses each derivation tree, checking features and using the derived feature sequences as the states of the internal nodes, where those derived states have as their 'output' argument the subtree below that node. In the transitive derivation, we slightly modify the trivial identity automaton to add case marking, as follows. In the transitive derivation (43), at the node labeled 2, the step taken by the identity transducer would be indicated this way:

$$\bullet(\text{=D V -EPP:V}(t_1),\text{D -EPP:D}(t_2)) \rightarrow \text{V -EPP:V,-EPP:D}(\bullet(t_1,t_2))$$

This indicates that at a node • where the left child has subtree t_1 and features =D V -EPP:V, and where the right node has subtree t_2 and features D -EPP:D, the resulting node should have features V -EPP:V,-EPP:D and subtree $\bullet(t_1,t_2)$. We mark case here by altering this step as follows:

(45) $\bullet(\text{=D V -EPP:V}(t_1),\ \text{D -EPP:D}(t_2)) \rightarrow \text{V -EPP:V,-EPP:D}(\bullet(\triangleleft(\text{H-},t_1),t_2))$

In effect, the symbol $\triangleleft$ inserted here just indicates that the tone is 'left adjoined' to the phrase at this point. That is, the spellout is unchanged except that an H- will be placed to the left of *le mamanu*.

In *tautala lelei*, the ergative is marked by a preceding /e/, and we can place that mark in the spellout too.[3] In the transitive derivation (43), at the node labeled 6, the step taken by the identity transducer would be indicated this way:

$$\bullet(\text{=D v}_{tr}\text{,-EPP:V}(t_1),\text{D}(t_2)) \rightarrow \text{v}_{tr}\text{,-EPP:V}(\bullet(t_1,t_2))$$

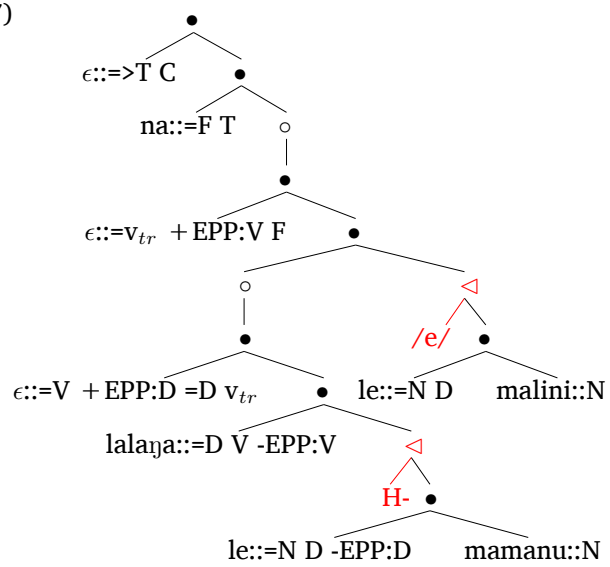
We interrupt the spellout here to place the case marker, in *tautala lelei* only:

[3]As noted in the text, an alternative idea is that the /e/ is a morpheme projecting a phrase in the syntactic structure. That alternative proposal could have been adopted here without affecting our main points.

(46) •(=D v_{tr},-EPP:V(t_1),D(t_2)) → v_{tr},-EPP:V(•($\triangleleft$(/e/, t_1), t_2))

With these rules we have case-marked the points in the derivation, with this result for *tautala lelei*:

(47)

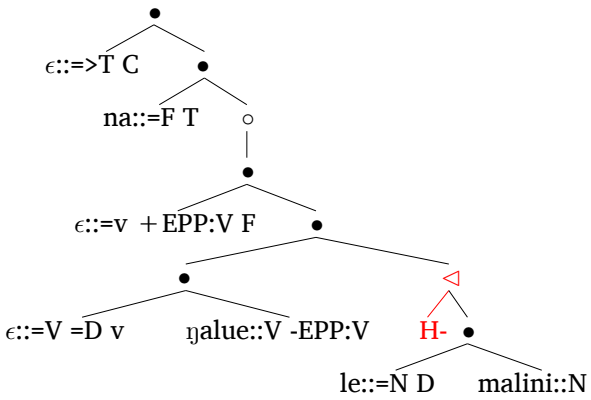


Spelling this out, we obtain the derived structure (11.a) for example (1b). The *tautala leaga* is similar, but lacks the ergative /e/. For our purposes, we do not need to resolve the interesting questions about what exactly explains the absence of the ergative marking in *tautala leaga* (see footnote 19).

For the intransitive (1b), we want to mark the absolutive in *tautala lelei*. In the derivation tree (43), the case marking needs to happen at node 3. By cancelling the features of lexical items, it can be seen that the left daughter there has the features =D v,-EPP:V, and the right daughter has the features D. So we can get the absolutive in this case with:

(48) •(=D v,-EPP:V(t_1), D(t_2)) → v,-EPP:V(•($\triangleleft$(H-, t_1), t_2)),

This rule transforms the derivation (43) to this one:



With no further change in spellout to strings, this gives us the derived structure (11b) for example (1a), as desired. In effect, all we have done in (46) and (48) is to represent the informal rules 'mark the ergative /e/' and 'mark the absolutive H', in a notation that succinctly defines the environments for the respective changes.

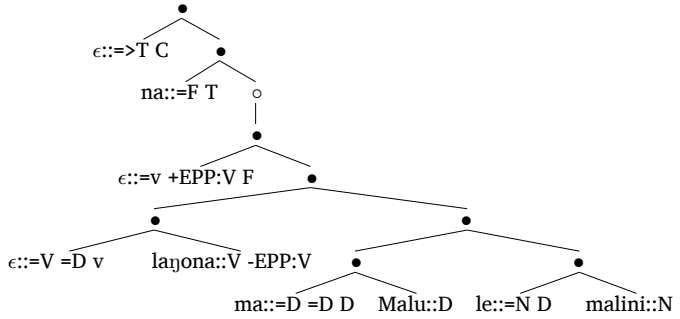
Note that with these rules and standard analyses of relativization, as desired, a high tone will not appear in relative clauses where an absolutive has been relativized. In standard analyses, the relativized position will be moving (on both promotion and operator-movement analyses), and so the syntactic features of the constituent there will differ. Since the syntactic features controlling spellout will differ, our rules will not

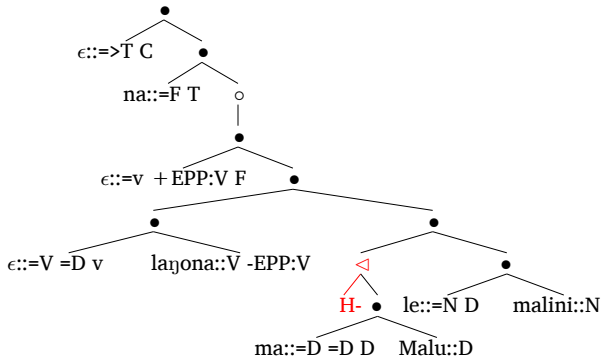
apply – no stipulation is needed. And that is the desired outcome; as mentioned earlier in Section 2, our data show no high tone marking there. Notice, on the other hand, that the absolutive case-marking rules (45) and (48) are similar, but not identical. They can be composed into a single function, but it might also be possible to adjust the representations so that both cases fall under one precise description; we leave that project to another place (and see Collins 2014).

For the coordinators, as noted in Section 5, we can either assume that the high tone is associated with the lexical items, or we can have a spellout rule:

(49) $\bullet$(=D D(t_1),D(t_2)) $\rightarrow$ D($\bullet(\triangleleft$(H-, t_1), t_2))

Note that case marking also applies to the transitive clauses in our examples of coordination, inserting the ergative case marker if the language is *tautala lelei*. But the effect of the coordination rule, by itself, is most often the insertion of H, transducing the tree on the left below to the tree on the right:

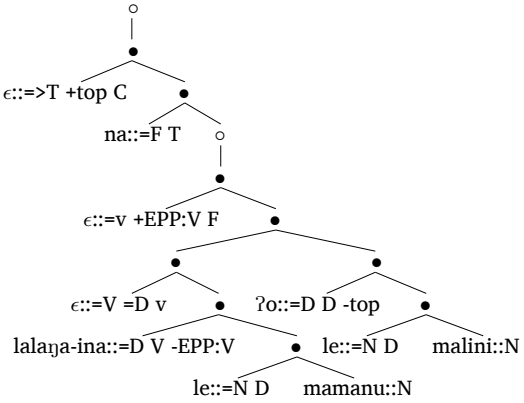


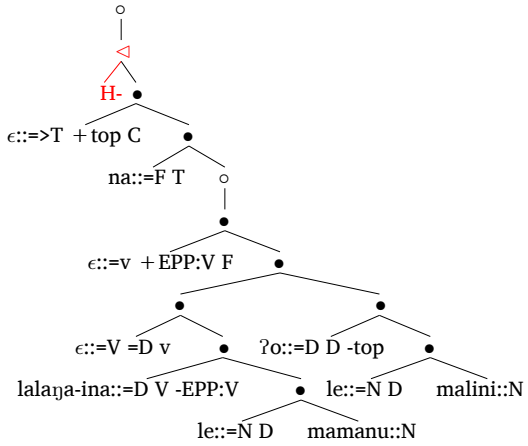


Similarly for the fronted argument in (19), we could have a spellout rule:

(50) $\circ$(+top C,-top(t_1)) $\rightarrow$ C($\circ(\triangleleft$(H-, t_1)))

The effect of the new rule, by itself, is the insertion of H- at the left edge of the clause, transducing the tree on the left below to the tree on the right:





While the previous steps define deterministic automata for inserting high tones, we see in Section 6 that the prosodic events defined by our rules can fail to surface for various reasons, reasons that are often not well understood. One way to model this situation is by making our model probabilistic, generalizing from deterministic to probabilistic automata. The formal properties of such probabilistic models have been well studied. We can, for example, replace the deterministic rule placing the absolutive H- (45) by a set of probabilistic rules that allow occasional L- or null marking in that position:

(45') •(=D V -EPP:V(t_1),D -EPP:D(t_2)) $\xrightarrow{0.98}$ V -EPP:V,-EPP:D(•($\lhd$(H-, t_1), t_2))
•(=D V -EPP:V(t_1),D -EPP:D(t_2)) $\xrightarrow{0.015}$ V -EPP:V,-EPP:D(•($\lhd$(L-, t_1), t_2))
•(=D V -EPP:V(t_1),D -EPP:D(t_2)) $\xrightarrow{0.005}$ V -EPP:V,-EPP:D(•(t_1, t_2))

The first rule applies to mark absolutive with H- 98% of the time. The second rule applies 1.5% of the time to mark absolutive with L. The last rule, with very low probability, leaves no mark at all. The rules for marking other positions can be modified similarly. For example, the coordination frequencies in Table 7 suggest that speaker f03 uses the rule marking coordinators with H approximately 85% of the time, with L otherwise (but probably null prosodic marking has non-zero probability too). The probabilities can vary some across speakers, as we have seen, and yet still provide reliable syntactic cues. With the probabilities evidenced in the data from any of our speakers, we will still find, for example, that the *tautala leaga* (8a) and (8b) will be distinguished with very high probability.

Composing syntax and post-syntax. The expressive power of MGs is quite well understood, and is exactly the same as several other independently proposed formalisms (Stabler, 2010). Furthermore, there are a number of quite general results about 'closure' properties of MGs; that is, results about how one language can be altered to produce another language in the same class. It is easy to establish that for any postsyntactic rules R like the ones described here, and any MG, the language that results from applying the rules to the derivations of the grammar L(R(MG)) is a language L(MG') defined directly by another grammar MG' without any post-syntactic rules. In fact, we can specify an expressively equivalent MG' in such a way that the derivations (and hence also the derived structures) of MG' are essentially identical to the derivations (and hence also the derived structures) R(MG).

The insertion rules we used above either left adjoined a constant to a defined set of phrases, or left adjoined a constant to a lexical item. We can 'compose' any such post-syntactic process into the grammar by modifying the lexicon of the grammar. We sketch how this can be done for one example; the others are similar. Consider the first post-syntactic rule from page viii, repeated here:

(45) •(=D V -EPP:V(t_1), D -EPP:D(t_2)) $\rightarrow$ V -EPP:V,-EPP:D(•($\lhd$(H, t_1), t_2))

This rule applies when a DP and a transitive V are combining; we see an example of its operation in the H-marked object of the derivation (47). This rule does not mark every DP, but only those DPs that are taken as an argument by a transitive verb. So we replace every lexical item for a transitive verb taking an object of category D by a lexical item that wants an object of category D', where D' is a category that appears nowhere else in the grammar; so in the lexicon above we replace

lalaŋa(-ina)::=D V -EPP:V 'weave', selects DP to form VP, then moves to EPP

by

lalaŋa(-ina)::=D' V -EPP:V 'weave', selects D'P to form VP, then moves to EPP

And for every lexical item of category D that remains in situ, we add a new item of category D' which is the same except that it selects a case marker of category A, where A is a category that occurs nowhere else in the grammar, and we add the H- as a lexical item of category A. So for the tiny grammar above, we add these lexical items:

le::=N =A D' (-EPP:D)	selects NP to form DP, optionally moves to +EPP:D
Malu::=A D'	name
ma::=D =D =A D'	coordinator 'and'
po'o::=D =D =A D'	coordinator 'or'
H-::A	the 'absolutive' high tone

Note that we do not add a new lexical item of category D' corresponding to the topicalizer ʔo, since that does not remain in situ. Also note that while the original lexical items for the coordinators, which are also in the new grammar, are recursive: They select DPs to make DPs. But the new lexical items for the coordinators of category D' are not recursive: They select DPs but they make D'. That D' is selected only by transitive verbs. Consequently, the resulting grammar, with these new lexical items, places the 'absolutive'

H- on in situ objects of transitive verbs in essentially the very same way the transducer rule (45) does. The structures defined by this grammar are isomorphic with the result of applying (45) to the structures defined by the original grammar.

After H- is placed into the sequence of morphemes by a post-syntactic process (one that has possibly been 'composed' into the syntax), we assume that additional post-syntactic processes apply: In particular, a tone-docking rule applies to dock the tone on the immediately preceding vowel. This too can be trivially accomplished with a finite state transducer (Kaplan & Kay, 1994; Heinz, 2011), and can similarly be 'composed into' a minimalist grammar to yield a grammar that parses the phonology directly. If Samoan were this simple, we would already have good parsers for it, parsers that use the tones in exactly the way we have prescribed. These parsers use H- to properly disambiguate the *tautala leaga* (8a) and (8b).

Note that all the rules can be composed into one, so that we compute, in one step, the value of dock(string(frontH(coordH(absH(erg(derivedTree)))))). That is, we can compute in one step the function that places the ergative, places the 'absolutive', places the coord H-, places the fronted H-, maps the tree to a phonological string, and docks the tone on the relevant phonological structure. Or we can separate some of these in the runtime parser, for example, separating the (un)docking rule from the rest. With any of these options, exactly the same strings with the same distributions of tonal events are predicted. And notice that even with these additional tones, these parsers still use H- to properly disambiguate the *tautala leaga* (8a) and (8b). So there are a range of mechanisms, adequate for the proposed structures, that the psycholinguist can consider when trying to understand the actual implementation of language production and recognition.

References

Collins, J. N. (2014). The distribution of unmarked cases in Samoan. In I. W. Arka & N. L. K. M. Indrawati (Eds.), *Papers from the 12th International Conference on Austronesian Linguistics, volume 2* (p. 93-110).

Collins, J. N. (2015). Diagnosing predicate fronting in Samoan. In U. Steindl et al. (Eds.), *Proceedings of the 32nd West Coast Conference on Formal Linguistics.* (Forthcoming)

Collins, J. N. (2016). Samoan predicate initial order and object positions. *Natural Language and Linguistic Theory*, *35*(1), 1–59.

Comon, H., Dauchet, M., Gilleron, R., Löding, C., Jacquemard, F., Lugiez, D., … Tommasi, M. (2007). *Tree automata techniques and applications.* (http://tata.gforge.inria.fr)

Graf, T. (2013). *Local and transderivational constraints in syntax and semantics* (Unpublished doctoral dissertation). UCLA.

Heinz, J. (2011). Computational phonology I: Foundations. *Language and Linguistics Compass*, *5*(4), 140-152.

Kaplan, R., & Kay, M. (1994). Regular models of phonological rule systems. *Computational Linguistics*, *20*, 331-378.

Kobele, G. M., Retoré, C., & Salvati, S. (2007). An automata-theoretic approach to minimalism. In J. Rogers & S. Kepser (Eds.), *Model theoretic syntax at 10, ESSLLI'07*.

Michaelis, J. (1998). Derivational minimalism is mildly context-sensitive. In *Proceedings, Logical Aspects of Computational Linguistics, LACL'98* (p. 179-198). NY: Springer.

Stabler, E. P. (2010). Computational perspectives on minimalism. In C. Boeckx (Ed.), *Oxford handbook of minimalism* (p. 617-641). Oxford: Oxford University Press.